Self-organized criticality in a network of interacting neurons

J D Cowan[1], J Neuman[2], W van Drongelen[3]

[1] Dept. of Mathematics, University of Chicago, 5734 S. University Ave., Chicago, IL 60637

[2] Dept. of Physics, University of Chicago, 5720 S. Ellis Ave., Chicago, IL, 60637

[3] Dept. of Pediatrics, University of Chicago, KCBD 900 E. 57th St., Chicago, IL., 60637

E-mail: `cowan@math.uchicago.edu`

Abstract. This paper contains an analysis of a simple neural network that exhibits self-organized criticality. Such criticality follows from the combination of a simple neural network with an excitatory feedback loop that generates bistability, in combination with an anti-Hebbian synapse in its input pathway. Using the methods of statistical field theory, we show how one can formulate the stochastic dynamics of such a network as the action of a path integral, which we then investigate using renormalization group methods. The results indicate that the network exhibits hysteresis in switching back and forward between its two stable states, each of which loses its stability at a saddle-node bifurcation. The renormalization group analysis shows that the fluctuations in the neighborhood of such bifurcations have the signature of directed percolation. Thus the network states undergo the neural analog of a phase transition in the universality class of directed percolation. The network replicates precisely the behavior of the original sand-pile model of Bak, Tang & Wiesenfeld.

1. Introduction

The idea of *self-organized criticality* (SOC) was introduced by Bak et al. (1988). Their paper immediately triggered an avalanche of papers on the topic, not the least of which was a connection with $1/f$- or scale-free noise. However it was not until another paper appeared, by Gil & Sornette (1996), which greatly clarified the dynamical prerequisites for achieving SOC, that a real understanding developed of the essential requirements for SOC: (1) an *order-parameter* equation for a dynamical system with a time-constant τ_o, with stable states separated by a threshold, (2) a *control-parameter* equation with a time-constant τ_c, and (3) a steady *driving force*. In Bak *et.al.*'s classic example, the sand-pile model, the order parameter is the rate of flow of sand grains down a sand-pile, the control parameter is the sand-pile's slope, and the driving force is a steady flow of grains of sand onto the top of the pile. Gil and Sornette showed that if $\tau_o \ll \tau_c$ then the resulting avalanches of sand down the pile would have a scale-free distribution, whereas if $\tau_o \gg \tau_c$ then the distribution would also exhibit one or more large avalanches.

In this paper we will analyze a neural network model which is in one-to-one correspondence with the Gil-Sornette SOC-model, and therefore also exhibits SOC.

1.1. Neural network dynamics

Consider first the mathematical representation of the dynamics of a neocortical slab comprising a single spatially homogeneous network of N excitatory neurons. Such neurons make transitions from a quiescent state q to an activated state a at the rate σ and back again to the quiescent state q at the rate α, as shown in figure 1.
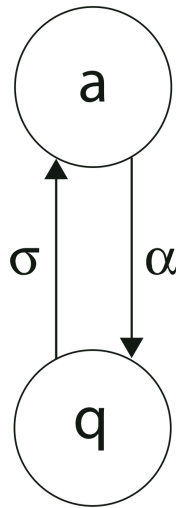

Figure 1. Neural state transitions

Let $P_n(t)$ be the probability that a fraction n/N is *active* at time t. Then the probabilistic dynamics of such a slab can be formulated as a master equation of the

form:

$$\frac{dP_n(t)}{dt} = \alpha[(n+1)P_{n+1} - nP_n] + (N-n+1)f[s(n-1)]P_{n-1}$$
$$- (N-n)f[s(n)]P_n \tag{1}$$

where α is the rate at which activated neurons become *quiescent*, and $s(n)$ is the total current or excitation driving each neuron in the population to fire at the rate $\sigma = f[s(n)]$. We assume in this simplified model that each neuron receives a signal weighted by w_0/N from each of N other neurons in the population, so that

$$s(n) = w_0 n + h \tag{2}$$

where h is an external current.

Equation 1 can be extended to the spatially inhomogeneous case. Let n_r/N_r be the fraction of active cells at time t in the rth population of N_r cells, and let $P[\mathbf{n}, t]$ be the probability of the configuration $\mathbf{n} = \{n_1, n_2, \cdots, n_r, \cdots, n_\Omega\}$ existing at time t. The extended master equation then takes the form

$$\frac{dP[\mathbf{n}, t]}{dt} = \alpha \sum_r [(n_r + 1)P[\mathbf{n}_{r+}, t] - n_r P[\mathbf{n}, t]]$$
$$+ \sum_r [(N_r - n_r + 1)f[s(n_r - 1)]P[\mathbf{n}_{r-}, t]$$
$$- (N_r - n_r)f[s(n_r)]P[\mathbf{n}, t]] \tag{3}$$

where $\mathbf{n}_{r\pm} = \{n_1, n_2, \cdots, n_r \pm 1, \cdots, n_\Omega\}$ and there are a total of Ω locally homogeneous populations.

Following Van Kampen (1981) equation 3 can be rewritten using the one-step operators

$$\mathcal{E}_r^\pm f(\mathbf{n}) = f(\mathbf{n}_r \pm 1) \tag{4}$$

i.e.,

$$\frac{dP[\mathbf{n}, t]}{dt} = \sum_r \left[\alpha(\mathcal{E}_r^+ - 1)n_r + (\mathcal{E}_r^- - 1)(N_r - n_r)f[s(n_r)] \right] P[\mathbf{n}, t] \tag{5}$$

Note that the total number of cells N_r in the rth population comprises n_r active cells and $N_r - n_r = q_r$ quiescent cells, so that

$$n_r + q_r = N_r \tag{6}$$

Thus equation 5 can be rewritten slightly in the form

$$\frac{dP[\mathbf{n}, t]}{dt} = \sum_r \left[\alpha(\mathcal{E}_r^+ - 1)n_r + (\mathcal{E}_r^- - 1)q_r f[s(n_r)] \right] P[\mathbf{n}, t] \tag{7}$$

1.2. Annihilation and creation operators

We now introduce Fock space *annihilation* and *creation* operators satisfying boson commutation rules

$$[a_r, a_s^\dagger] = [q_r, q_s^\dagger] = \delta(r - s)$$
$$[a_r, a_s] = [a_r^\dagger, a\dagger_s] = 0$$
$$[q_r, q_s] = [q_r^\dagger, q_s^\dagger] = 0 \tag{8}$$

such that given neural vectors $|n_r\rangle$, $|m_{r'}\rangle$

$$a_r^\dagger|n_r\rangle = |n_r + 1\rangle, a_r|n_r\rangle = n_r|n_r - 1\rangle$$
$$q_r^\dagger|m_r\rangle = |m_r + 1\rangle, q_r|m_r\rangle = m_r|m_r - 1\rangle \tag{9}$$

where

$$n_r + m_r = 1, |n_r\rangle = (a^\dagger)_r^n|0_r\rangle, |m_r\rangle = (q_r^\dagger)_r^m|0_r\rangle \tag{10}$$

and $|0_r\rangle$ is a fiducial or vacuum (empty) state such that

$$a_r^\dagger|0_r\rangle = |n_r = 1_r\rangle, a_r|0_r\rangle = 0$$
$$q_r^\dagger|0_r\rangle = |m_r = 1_r\rangle, q_r|0_r\rangle = 0 \tag{11}$$

and there exists a *dual* vector $\langle 0|$ such that

$$\langle 0_r|a_r = \langle n_r = 1_r|, \langle 0_r|a_r^\dagger = 0,$$
$$\langle 0_r|q_r = \langle m_r = 1_r|, \langle 0_r|q_r^\dagger = 0, \tag{12}$$

It follows from equation 9 that

$$a_r^\dagger a_r|n_r\rangle = n_r|n_r\rangle$$

i.e. n_r is the *eigenvalue* of the operator $a_r^\dagger a_r$, which is therefore referred to as a *number* or *number density* operator.

1.3. Bosonizing the master equation

The master equation can now be *bosonized* by replacing the van Kampen one-step operators by bosonic equivalents, i.e.

$$(\mathcal{E}_r^+ - 1) = (q_r^\dagger - a_r^\dagger)a_r$$
$$(\mathcal{E}_r^- - 1) = (a_r^\dagger - q_r^\dagger)q_r \tag{13}$$

so eqn 7 becomes:

$$\frac{dP[\mathbf{n}, t]}{dt} = \sum_r \left[\alpha(q_r^\dagger - a_r^\dagger)a_r + (a_r^\dagger - q_r^\dagger)q_r f[s(a_r^\dagger a_r)] \right] P[\mathbf{n}, t] \tag{14}$$

Let $|P(t)\rangle$ be a probability state vector satisfying

$$|P(t)\rangle = \sum_\mathbf{n} P[\mathbf{n}, t]|\mathbf{n}\rangle \tag{15}$$

Then equation 7 can be written in the form

$$\frac{d}{dt}|P(t)\rangle = \sum_r \left[\alpha(q_r^\dagger - a_r^\dagger)a_r + (a_r^\dagger - q_r^\dagger)q_r f[s(a_r^\dagger a_r)]\right]|P(t)\rangle \qquad (16)$$

or formally as

$$\frac{d}{dt}|P(t)\rangle = -\hat{H}|P(t)\rangle \qquad (17)$$

where

$$-\hat{H} = \sum_r \left[\alpha(q_r^\dagger - a_r^\dagger)a_r + (a_r^\dagger - q_r^\dagger)q_r f[s(a_r^\dagger a_r)]\right] \qquad (18)$$

is the *quasi-Hamiltonian* operator for the Markov process represented in equation 1.

1.4. From bosons to coherent states

Equation 18 is a linear operator equation with formal solution

$$|P(t)\rangle = \exp[-\hat{H}(t - t_0)|P(t_0)\rangle$$

We need to re-express this solution in terms of numbers rather than operators. This can be achieved by introducing *coherent states*. These were introduced by Schrödinger (1926) and first used extensively in coherent optics by Glauber (1963). We therefore introduce such states $|\phi_r\rangle$ in the form

$$|\phi_r\rangle = \exp[-\frac{1}{2}\varphi_r^\star \varphi_r + \varphi_r a_r^\dagger]|0_r\rangle \qquad (19)$$

where φ_r is the right eigenvalue of a_r, i.e. $a_r|\phi_r\rangle = \varphi_r|\phi_r\rangle$. There is also a coherent state representation of q_r in the form $|\theta_r\rangle$ such that the right eigenvalue of q_r is ϑ_r, i.e. $q_r|\theta_r\rangle = \vartheta_r|\theta_r\rangle$. In similar fashion $\langle\phi_r|a_r^\dagger = \langle\phi_r|\tilde{\varphi}_r$ where $\tilde{\varphi}_r$, the complex conjugate of φ, is the left eigenvalue of $a_r^\dagger$, and similarly $\langle\theta_r|q_r^\dagger = \langle\theta_r|\tilde{\vartheta}_r$, i.e. $\tilde{\vartheta}_r$ is the left eigenvalue of $q_r^\dagger$. It follows that

$$\langle\phi_r|a_r^\dagger a_r|\phi_r\rangle = \langle\phi_r|\tilde{\varphi}_r \varphi_r|\phi_r\rangle = \tilde{\varphi}_r \varphi_r \qquad (20)$$

All this suggests that the operator quasi-Hamiltonian has a coherent state representation in the form

$$-\mathcal{H} = \sum_r \left[\alpha(\tilde{\vartheta}_r - \tilde{\varphi}_r)\varphi_r + (\tilde{\varphi}_r - \tilde{\vartheta}_r)\vartheta_r f[s(\tilde{\varphi}_r \varphi_r)]\right] \qquad (21)$$

Note in passing that operator products involving powers of the number operator $a_r^\dagger a_r$ must first be *normal ordered*, i.e. all creation operators $a_r^\dagger$ must preceed the annihilation operators a_r, before coherent states can be introduced. For example the normal order form of $\exp[a^\dagger a]$ written as $: \exp[a^\dagger a] :$ is expanded as

$$: \exp[a^\dagger a] : \ = 1 + (a^\dagger a) + \frac{1}{2!}(a^\dagger a)^2 + \cdots$$

$$= 1 + (a^\dagger a) + \frac{1}{2!}(a^\dagger a + a^{\dagger 2}a^2) + \cdots$$

$$= \sum_{l=0}^{\infty}\frac{1}{l!}\sum_{j=0}^{l}s_{l,j}a^{\dagger j}a^j \qquad (22)$$

where the $s_{l,j}$ are Stirling numbers of the second kind. It follows that $:\exp[a^\dagger a]:$ can be written as

$$:\exp[a^\dagger a]:= \sum_{k=0} h_k a^{\dagger k} a^k \tag{23}$$

where $h_k = \sum_l s_{l,k}$.

The final preliminary step in this formulation is to take the *continuum limit* of the expression for $\mathcal{H}$ in equation 21, so that

$$-\mathcal{H} = \int\int d^d x dt \left[\alpha(\tilde{\vartheta} - \tilde{\varphi})\varphi + (\tilde{\varphi} - \tilde{\vartheta})\vartheta f[s(\tilde{\varphi}\varphi + \varphi)]\right] \tag{24}$$

where $\varphi_r \to \varphi(\mathbf{x}, t) \equiv \varphi$ etc., and the conjugate coherent state $\tilde{\varphi}$ has been shifted to $1 + \tilde{\varphi}$.

1.5. Dimensions and the density representation

Before proceeding further we need to assign a dimension to each variable in equation 24. To do so we use a modified version of the convention used in particle physics so that $[x] = L^{-1}$, $[t] = L^{-2}$ where L is the length scale used, whence $[x^2/t] = L^0$. {This generates a scaling found in Markov random walks and related processes such as stochastic neural activity.} Then $[\alpha] = L^2$, $[\varphi] = L^d$, $[\tilde{\varphi}] = L^0$, $[\tilde{\varphi}\varphi] = L^d$, $[f[s]] = [\alpha] = L^2$. This last value of $[f[s]]$ implies that the input current function $s(\tilde{\varphi}\varphi + \varphi) = s(I) = kI$ where the constant k has the dimensions of inverse current. The net effect of such a choice leads to the required result that $[\mathcal{H}] = 0$.

To emphasize this choice we further transform the coherent-state quasi-Hamiltonian by introducing the *density representation*:

$$\tilde{\varphi} \to \exp[\tilde{n}] - 1, \varphi \to n[\exp[-\tilde{[n]}]]$$
$$\tilde{\vartheta} \to \exp[\tilde{p}] - 1, \vartheta \to p[\exp[-\tilde{[p]}]] \tag{25}$$

so that equation 24 transforms into

$$-\mathcal{H} = \int\int d^d x dt \left[\alpha(\exp(\tilde{p} - \tilde{n}) - 1)n + (\exp(\tilde{n} - \tilde{p}) - 1)pf[s(n)]\right] \tag{26}$$

Note that in the continuum limit the input current function $s(n) = s(I) = kI = k(w \star n + h)$ where $\star$ is the spatial convolution operator, i.e. $w \star n = \int d^d \mathbf{x}' w(\mathbf{x} - \mathbf{x}')n(\mathbf{x}', t)$.

1.6. From the quasi-Hamiltonian to a neural Path Integral

Using standard methods (Doi 1976, Peliti 1985) Buice & Cowan (2007) incorporated the quasi-Hamiltonian into the action of a Wiener path integral. This action takes the form:

$$S(n) = \int\int d^d x dt \, [\tilde{n}\partial_t n + \tilde{p}\partial_t p +$$
$$\alpha(1 - \exp(-(\tilde{n} - \tilde{p}))n - (\exp(\tilde{n} - \tilde{p}) - 1)pf[s(n)]] \tag{27}$$

The utility of this action is that it is part of the exponent of the moment generating functional for the statistical moments of the probability density $P[\mathbf{n}, t]$.

At an extremum

$$\left.\frac{\delta S(n)}{\delta \tilde{n}}\right|_{\tilde{n}=0} = \left.\frac{\delta S(n)}{\delta \tilde{p}}\right|_{\tilde{p}=0} = 0 \tag{28}$$

which leads to the mean-field equations

$$\partial_t n + \alpha n - pf[s(n)] = 0, \quad \partial_t p - \alpha n + pf[s(n)] = 0 \tag{29}$$

whence

$$n + p = \rho \tag{30}$$

where ρ is the (constant) packing density of excitatory neurons in the population. This is a mean-field result, so we replace p by $\rho - n_{cl}$ in equation 27, so as to generate the mean-field Wilson-Cowan equation (Wilson & Cowan 1973) for a single spatially organized population from the first variation of the action given in equation 27 and (in principle), all higher moments, and (finally) rewrite equation 27 in the form:

$$S(n) = \int\int d^d x dt \, [\tilde{n}\partial_t n +$$
$$\alpha(1 - \exp(-\tilde{n}))n - (\exp(\tilde{n}) - 1)(\rho - n_{cl})f[s(n)]] \tag{31}$$

1.7. Renormalizing the path integral

We expand n about its mean value $\langle n \rangle = n_{cl}$, which satisfies equation 28 in the form:

$$\partial_t n_{cl} = -\alpha n_{cl} + (\rho - n_{cl})f[s(n_{cl})] \tag{32}$$

where

$$s(n_{cl}) = k(w \star n_{cl} + h_{cl}) \tag{33}$$

Thus $n \to n + n_{cl}, \tilde{n} \to \tilde{n}$, since $\tilde{n}_{cl} = 0$. So $s(n) \to s(n + n_{cl})$ and $f[s(n)] \to f[s(n + n_{cl})]$. If follows that $s(n + n_{cl}) = k(w \star (n + n_{cl}) + (h + h_{cl})) = k(w \star n_{cl} + h_{cl}) + k(w \star n + h)) = s(n_{cl}) + s(n)$, and therefore $f[s(n)] = f[s(n_{cl}) + s(n)]$. We next expand $f[s(n)]$ in a Taylor expansion about the mean-field value n_{cl}, noting that from equation A.1, $s(n) = k(w \star n + h)) = k(Ln + h)$. In what immediately follows we assume that the external stimulus $h(\mathbf{x}, t) = 0$. It follows that:

$$f[s(n)] = f[kL(n_{cl} + n)] = f[kLn_{cl}] + f^{(1)}[kLn_{cl}]kLn$$
$$+ \frac{1}{2}f^{(2)}[kLn_{cl}](kLn)^2 + \cdots \tag{34}$$

However because of normal ordering, equation 34 leads to the expression:

$$f[s(n)] = \sum_m g_m(kLn)^m, \quad \text{where} \quad g_m = \sum_{l=0}^m \frac{f^{(l)}}{l!}s_{l,m} \tag{35}$$

Since the leading terms of g_m are proportional to $f^{(m)}$, and given the assumed form for $f[s(n)]$ to be such that $f^{(1)} > 0$ and $f^{(2)} < 0$, then $g_m > 0$ for m odd, and $g_m < 0$ for m even.

We also expand the functions $\exp(\pm\tilde{n})$. The resulting action $S(n)$ takes the form:

$$S(n) = \int\int d^dx dt \left[\tilde{n}(\partial_t + \alpha - (\rho - n_{cl})g_1 kL)n\right.$$

$$-\frac{\tilde{n}^2}{2}(\alpha + (\rho - n_{cl})g_1 kL)n + \tilde{n}((\rho - n_{cl})g_2(kL)^2)n^2$$

$$\left.+\frac{\tilde{n}^2}{2}((\rho - n_{cl})g_2(kL)^2)n^2 + \cdots\right] \tag{36}$$

It follows from the appendix that w_2 is small compared to w_0, so that in most expressions the terms proportional to $\nabla^{2m}n^m$ can be neglected. However this is not always the case for $m = 1$. Thus the first term can be written approximately as $\tilde{n}(\partial_t + \alpha - (\rho - n_{cl})g_1 k(w_0 + \frac{1}{2!}w_2\nabla^2))n = \tilde{n}(\partial_t + \mu - D\nabla^2)n$ where $\mu = \alpha - (\rho - n_{cl})g_1 kw_0$ and $D = \frac{1}{2}(\rho - n_{cl})g_1 kw_2$. So the expression for the action is now reduced to the form:

$$S(n) = \int\int d^dx dt \left[\tilde{n}(\partial_t + \mu - D\nabla^2)n - \tilde{n}^2 G_1 n\right.$$

$$\left.+\tilde{n}G_2 n^2 + \frac{1}{2}\tilde{n}^2 G_2 n^2 + \cdots\right] \tag{37}$$

where $G_1 = 1/2(\alpha + (\rho - n_{cl})g_1 kw_0)$, and $G_2 = (\rho - n_{cl})|g_2|k^2 w_0^2$. We need to demonstrate that the last term in $S(n)$, i.e., $\frac{1}{2}\tilde{n}^2 G_2 n^2$, and all other terms, are *irrelevant* in the sense of the renormalization group.

The renormalization group [RG] analysis is carried out via dimensional analysis. It can be shown that all the terms in $S(n)$ are zero-dimensional when integrated over d-dimensional space and over time, i.e., $[d^dx dt] = L^{-(d+2)}$ and [any term in the integrand] $= L^{d+2}$. However, as it stands $[n] = L^d$, but $[\tilde{n}] = L^0$, so that $[\tilde{n}n] = L^{0+d} = L^d$. This is not suitable for the scaling analysis implemented in the RG process. We therefore introduce a new *scaling*,

$$\tilde{s} = \sqrt{\frac{G_1}{G_2}}\tilde{n}, \quad s = \sqrt{\frac{G_2}{G_1}}n \tag{38}$$

such that $\tilde{s}s = \tilde{n}n$ where $[G_2/G_1] = L^{-d}$. The effect of this scaling is that both $\tilde{s}$ and s have dimension $L^{d/2}$. Let

$$\sqrt{G_1 G_2} = u, \quad G_2 = 2\tau \tag{39}$$

The net effect of this scaling transformation is that

$$S(s) = \int\int d^dx dt \left[\tilde{s}(\partial_t + \mu - D\nabla^2)s + u\tilde{s}(s - \tilde{s})s + \tau\tilde{s}^2 s^2 + \cdots\right] \tag{40}$$

But $[\tau/u] = L^{-d/2}$ and therefore scales to zero as $L \to \infty$ under subsequent RG transforms. So asymptotically the terms $\tau\tilde{s}^2 s^2 + \cdots$ become *irrelevant* in the RG sense. So finally

$$S(s) = \int\int d^dx dt \left[\tilde{s}(\partial_t + \mu - D\nabla^2)s + u\tilde{s}(s - \tilde{s})s\right] \tag{41}$$

is the *renormalized action* of the large-scale neural activity of a single neural population.

1.8. Directed Percolation

This action is well-known: it is called Reggeon Field Theory, and is found in directed percolation [DP] in random graphs, in contact processes, in high-energy nuclear physics, in bacterial colonies, all of which exhibit the characteristic properties of what is called a *universality class*, i.e., it is a phase transition with a universal scaling of important statistical exponents. It also shows up in branching and annihilating random walks, catalytic reactions, and interacting particles. Thus we have mapped the mathematics of large-scale neural activity in a single homogeneous neural population into a percolation problem in random graphs, or equivalently into a branching and annihilating random walk. A first version of this work was presented in Buice & Cowan (2007). A more extensive paper with many applications to neuroscience was presented in Buice & Cowan (2009).

Here we note that there is an *upper critical dimension* at which directed percolation crosses over to mean-field behavior. This upper critical dimension is $d = 4$. What is the dimension of the neocortex? To answer this question we look at the number of synapses per neuron in the neocortex. Using estimates provided by Stevens (1989), this number is about 4×10^3. Assuming the number of synapses in a terminal axonal arbor to be about 50, the number of neural neighbors per neuron is about 80, so the effective dimensionality of a neocortical hyper lattice is about $d = 40$. Thus the critical exponents characterizing the neural phase transition are the $d = 4$ exponents of directed percolation. These have been calculated by Abarbanel & Broznan (1974), Abarbanel et al. (1976), and Amati et al. (1976), and appear in the linear response of the neocortical model to an impulsive stimulus, known to mathematicians as the *Green's function* and to physicists as the *propagator*. This takes the form

$$G(x - x', t - t') \propto \begin{cases} (t - t')^{-2} \exp\left[-\frac{x - x')^2}{4(t - t')} - \mu(t - t')\right], & \mu > 0 \\ (t - t')^{-2} \exp\left[-\frac{x - x')^2}{4(t - t')})\right], & \mu = 0 \\ \mu^2 \Theta\left[\sqrt{|\mu|}(t - t') - |x - x'|\right], & \mu < 0 \end{cases} \quad (42)$$

where the cases $\mu > 0, \mu = 0$, and $\mu < 0$ correspond, respectively, to the sub-critical, critical, and super-critical propagators. They correspond, respectively, to solutions of the cable equation, the diffusion equation, and a nonlinear wave-equation. The critical propagator is thus the diffusion limit of Brownian motion. It turns out that there is a great deal of data supporting the hypothesis that the mean-field propagator of DP correctly describes the essential features of large-scale neocortical activity on many spatio-temporal scales. [See Burns (1951), Lampl et al. (1999), Nauhaus et al. (2009).] In addition data on the statistical structure of large-scale activity recorded in cortical slices by Beggs & Plenz (2003) supports the hypothesis. In particular, the avalanche-size distribution of spontaneous activity in cortical slices fits the DP hypothesis.

The analysis can be extended to deal with a neural network comprising both excitatory and inhibitory neurons. However in this paper we describe how to incorporate synaptic plasticity into a network comprising only excitatory neurons, as a mechanism

that *tunes* the network so that it automatically reaches the critical point of the DP phase transition, thus exhibiting *self-organized criticality.*

2. Incorporating synaptic plasticity

Consider first a single excitatory population model with a fixed recurrent excitatory synapse w_E and an input H through an excitatory modifiable synapse w_H. The mean-
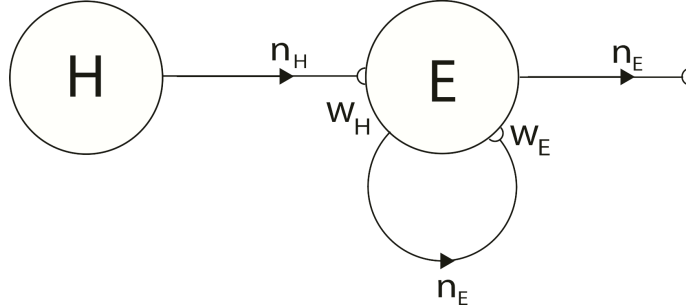

Figure 2. A recurrent excitatory network

field neural equation for this is a version of equation 32, i.e.

$$\frac{dn_E}{dt} = -\alpha_E n_E + (1 - n_E)\sigma_E\left[s_E(I_E)\right] \tag{43}$$

where

$$s_E(I_E) = I_E/I_{RH,E}, \quad I_E = w_E \star n_E + w_H \star n_H \tag{44}$$

and the synapse w_H is *modifiable* with

$$\frac{dw_H}{dt} = -g_E \langle (n_E - \rho_{E,0} - \rho_{E,S} w_H) n_H \rangle_t \tag{45}$$

where $\rho_{E,0}$ is a constant neural activity, $\rho_{E,S}$ is a constant (Vogels et al. 2011), and g_E is a state-dependent rate function.

This is also a mean-field equation in which the synaptic weight w_H is depressed by an *anti-Hebbian* mechanism, and potentiated by the input activity n_H. The problem is to write an action that incorporates these equations. Note that the time scale of the growth and decay of neural activity is set by the constant α_E, whereas that of the growth and decay of synaptic plasticity is set by g_E. Thus the ratio α_E/g_E is an important parameter.

3. Deriving an action for synaptic plasticity

The first problem is to develop an action for the modifiable synapse w_H. In order to do so we first note from equation A.2 that w_H scales with b_H, the synaptic conductance or *weight*, and we can write an approximation to eqn. 45 in the form:

$$\frac{db_H}{dt} = -g_E \left(n_E - \rho_{E,0} - \rho_{E,S} k_H b_H\right) n_H \tag{46}$$

We next reformulate the changes in b_H as a Markov process with discrete states in continuous time. We therefore assume that b_H is *quantized* in units Δ of synaptic weight, and similarly for b_E. Thus

$$b_E = m_E \Delta, \quad b_H = m_H \Delta \tag{47}$$

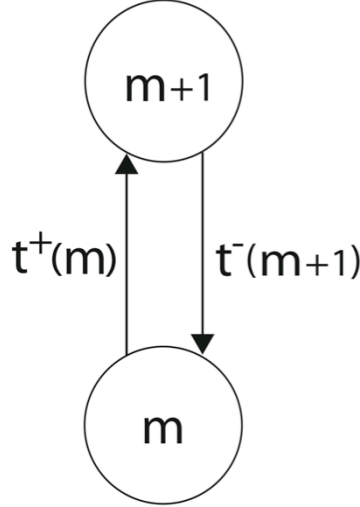and we look at the Markov process represented in figure 3:

Figure 3. Synaptic state transitions

and we write a *master equation* for this process in the form

$$
\begin{aligned}
\frac{dP(m_H, t)}{dt} &= t^-(m_H + 1)P(m_H + 1, t) - t^-(m_H)P(m_H, t) \\
&\quad + t^+(m_H - 1)P(m_H - 1, t) - t^+(m_H)P(m_H, t)
\end{aligned}
\tag{48}
$$

where $P(m, t)$ is the probability that the synaptic weight $m_H = m$ at time t. and

$$
\begin{aligned}
t^+ &= g_E(\rho_{E,0} + \rho_{E,S} k_H m_H) n_H \\
t^- &= g_E n_E n_H
\end{aligned}
\tag{49}
$$

are the transition rates for the excitatory synapse m_H.

3.1. The van Kampen ladder operators

We again introduce the van Kampen ladder operators

$$E_m^{\pm} m = m \pm 1 \tag{50}$$

so that the master equation can be rewritten as:

$$\frac{dP(m_H, t)}{dt} = \left[t^-(E_{m_H}^+ - 1) + t^+(E_{m_H}^- - 1) \right] P(m_H, t) \tag{51}$$

An examination of eqns. 48-50 indicates that only the transition rate t^+ contains a term proportional to m_H. We will utilize this property in what follows.

3.2. From the Master Equation to the Action

We now introduce bosonic annihilation and creation operators for m_H. Let such operators be denoted by $s^\dagger$ and s respectively, and let $|m\rangle$ be a column vector representing the synaptic weight m such that

$$s^\dagger|m\rangle = |m+1\rangle, \quad s|m\rangle = m|m-1\rangle \tag{52}$$

Note that $s^\dagger$ and s satisfy the same commutation rules and equations that we introduced earlier, so that

$$E_m^+ - 1 \to s^\dagger - s^\dagger s = s^\dagger(1-s)$$
$$E_m^- - 1 \to s - s^\dagger s = (1-s^\dagger)s \tag{53}$$

So we consider the equation:

$$\frac{\partial|\hat{m}\rangle}{\partial t} = (E_m^- - 1)|m\rangle = (1-s^\dagger)s|m\rangle \tag{54}$$

We further note that $(1-s^\dagger)s$ is *normal ordered*. We therefore shift $s^\dagger$ and s,

$$s^\dagger \to 1 + \tilde{s}, s \to s \tag{55}$$

where $\tilde{s}$ and s are now *coherent states*, and introduce the *density representation*

$$\tilde{s} \to \exp(\tilde{m}) - 1, \quad s \to m\exp(-\tilde{m}) \tag{56}$$

we find

$$\partial_t|m\rangle = -\tilde{s}s|m\rangle$$
$$= -(\exp(\tilde{m})-1)m\exp(-\tilde{m})|m\rangle$$
$$= -(1-\exp(-\tilde{m}))m|m\rangle \tag{57}$$

Eqn. 21 tells us that the action S_m must contain a term of the form

$$-g_E\rho_{E,S}k_H n_H(1-\exp(-\tilde{m}_H))m_H,$$

a *source* term

$$-g_E\rho_{E,0}n_H\tilde{m}_H,$$

and an *interaction* term of the form

$$+g_E n_E n_H\tilde{m}_H,$$

leading to an action of the form:

$$S(m_H) = \int\int d^dx dt\,[\tilde{m}_H\partial_t m_H - g_E\rho_{E,S}k_H n_H(1-\exp(-\tilde{m}_H))m_H$$
$$-\tilde{m}_H g_E\rho_{E,0}n_H + \tilde{m}_H g_E n_E n_H] \tag{58}$$

Using variational techniques, we can derive the mean-field equation [eqn 45] from the condition:

$$\left.\frac{\delta S(\tilde{m}_H)}{\delta\tilde{m}_H}\right|_{\tilde{m}_H=0} = 0 \tag{59}$$

For then we obtain the equation:

$$\frac{dm_H}{dt} = -g_E\left(n_E - \rho_{E,0} - \rho_{E,S}k_H m_H\right)n_H \tag{60}$$

i.e., eqn. 46.

3.3. Renormalizing the synaptic plasticity action

We now proceed to renormalize the action $S(m_H)$ just as we renormalized $S(n_E)$. We therefore expand the exponential term in equation 58 and rewrite $S(m_H)$ in the form:

$$S(m_H) = \int \int d^d x dt \left[\tilde{m}_H \partial_t m_H - H_1 \tilde{m}_H m_H n_H + \frac{1}{2} H_1 \tilde{m}_H^2 m_H n_H \right.$$
$$\left. - H_2 \tilde{m}_H n_H + \tilde{m}_H g_E n_E n_H \right] \tag{61}$$

where $H_1 = g_E \rho_{E,S} k_H$, and $H_2 = g_E \rho_{E,0}$.
We now introduce the scaling

$$\tilde{s}_H = \sqrt{\frac{H_2}{H_1}} \tilde{m}_H, \quad s_H = \sqrt{\frac{H_1}{H_2}} m_H \tag{62}$$

such that $\tilde{s}_H s_H - \tilde{m}_H m_H$, and $[H_1/H_2] = L^{-d}$. This scaling is analogous to the scaling of n and $\tilde{n}$ which we carried out earlier for neural activities. As before the effect of this scaling is that both $\tilde{s}_H$ and s_H have dimension $L^{d/2}$.
Let

$$\sqrt{H_1 H_2} = u_H, \quad H_1 = 2\tau_H \tag{63}$$

and recall that equation 38 scales n_E to $\sqrt{G_1/G_2} s_E$.

Following the procedure outlined earlier we can calculate which terms in the transformed action $S(s_H)$ become irrelevant under scaling transformations. The resulting renormalized synaptic plasticity action takes the form:

$$S(s_H) = \int \int d^d x dt \left[\tilde{s}_H \partial_t s_H - u_H \tilde{s}_H n_H \right] \tag{64}$$

4. Combining the actions

It follows from this formulation that the full action for the coupled system of equations for the evolution of n_E and m_H can be obtained simply by adding the actions $S(n_E)$ and $S(m_H)$ together. The combined action therefore takes the form:

$$S(n_E, m_H) = \int \int d^d x dt \left[\tilde{n}_E \partial_t n_E + \alpha (1 - \exp(-\tilde{n}_E)) n_E \right.$$
$$- (\exp(\tilde{n}_E) - 1)(\rho - n_{E,cl}) f[s(n_E)] + \tilde{m}_H \partial_t m_H$$
$$- g_E \rho_{E,S} k_H n_H (1 - \exp(-\tilde{m}_H)) m_H$$
$$\left. - \tilde{m}_H g_E \rho_{E,0} n_H + \tilde{m}_H g_E n_E n_H \right] \tag{65}$$

4.1. A simulation of the behavior of the combined mean-field equations

The first variation of equation 65 generates the mean-field equations for n_E and m_H in the form:

$$\frac{dn_E}{dt} = -\alpha_E n_E + (1 - n_E)\sigma_E [s_E(I_E)]$$
$$\frac{dm_H}{dt} = -g_E (n_E - \rho_{E,0} - \rho_{E,S} k_H m_H) n_H \tag{66}$$

where

$$s_E(I_E) = k_E(m_E \star n_E + m_H \star n_H) \tag{67}$$

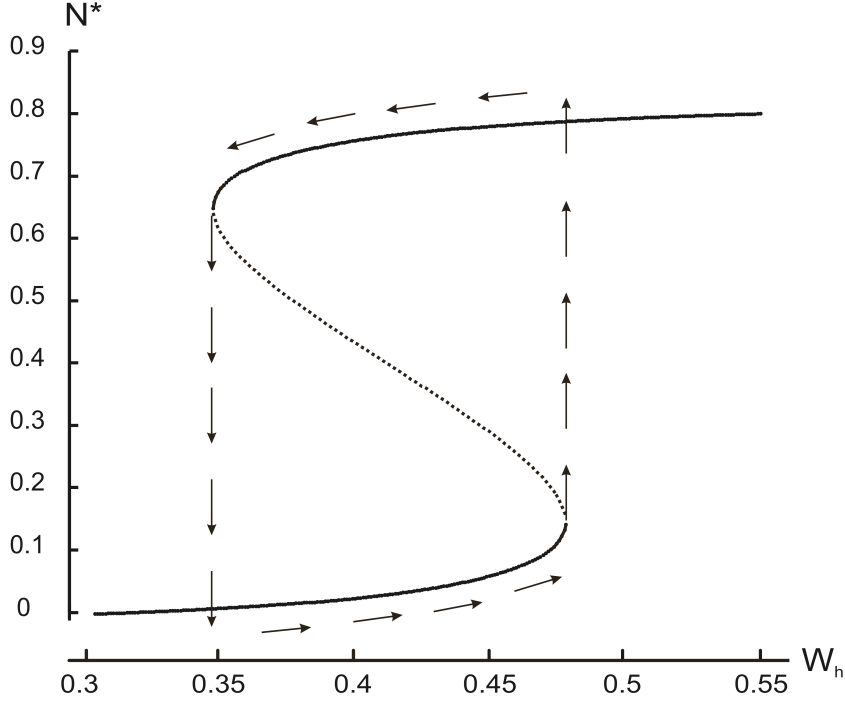These equations can be simulated. The results are shown in figure 4. It will be seen that

Figure 4. Neural state transitions between a ground state and an excited state. Parameter values: $m_E = 3, n_H = 3; \alpha = 0.2$. $N^\star$ is the fixed-point value of n_E, and W_H is the magnitude of the anti-Hebbian synapse in the input path.

in the "ground-state" of low values of $N^\star = n_E^\star$ the synaptic weight w_H (proportional to m_H), increases until it reaches the critical point at a saddle-node bifurcation, at which point $N^\star$ becomes unstable and the system switches to the "excited-state". But then the anti-Hebbian term in the synaptic plasticity dynamics kicks in, and W_H declines until the excited-state fixed-point becomes unstable at the upper critical point, also at a saddle-node, and switches back to the ground-state fixed point, following which the hysteresis cycle starts over. This is a exact representation of the sand-pile model's behavior. This is another representation of SOC in a neural network. The reader should compare this with the synaptic mechanisms for achieving SOC described in Levina et al. (2007) and in Millman et al. (2010).

4.2. Renormalizing the combined action

To renormalize the combined action we follow the same procedure as before and expand the exponential functions $\exp[\pm\tilde{n}_E]$, $\exp[\pm\tilde{m}_H]$, and $f[s(n_E)]$. Note that after normal ordering and collecting terms, $f[s(n_E)]$ can be expanded in the extended form

$f[s(n_E)] = \sum_m g_m (k_E(L_E n_E + L_H n_H))^m$. After scaling and dimensional analysis, the renormalized combined action takes the form:

$$S(s_E, s_H) = \int \int d^d x dt \left[\tilde{s}_E(\partial_t + \mu_E - D_E \nabla^2)s_E + u_E \tilde{s}_E(s_E - \tilde{s}_E)s_E \right.$$
$$\left. - v_E \tilde{s}_E n_H + \tilde{s}_H \partial_t s_H - u_H \tilde{s}_H n_H \right] \tag{68}$$

where u_E, u_H and v_E are renormalized constants. It would appear that apart from the source term $v_E \tilde{s}_E n_H$ the coupled action is just the sum of the two uncoupled renormalized actions. This is indeed the case! All the addition terms which appear in the current function become irrelevant under renormalization, and do not effect the renormalized action representing n_E. Similarly for m_H. It follows that fluctuations in the activity n_E in the neighborhood of the critical point $\mu_E = 0$, i.e. those in the fluctuation-driven regime, should be essentially those of directed percolation.

5. Inhibitory synapses

In case w_H is an *inhibitory* synapse the transitions $t^+(m)$ and $t^-(m)$ are reversed. so that:

$$t^+ = g_E n_E n_H$$
$$t^- = g_E(\rho_{E,0} + \rho_{E,S} k_H m_H)n_H \tag{69}$$

Thus now only the transition rate t^- contains a term proportional to m_H. The effect of this is that the action for S_m leads to the mean-field equation:

$$\frac{dm_H}{dt} = g_E \left(n_E - \rho_{E,0} - \rho_{E,S} k_H m_H \right) n_H \tag{70}$$

at an inhibitory synapse m_H.

Thus inhibitory feedforward synapses are Hebbian with stimulus dependent depression, whereas excitatory feedforward synapses are anti-Hebbian with stimulus dependent potentiation, and we have now formulated actions for feedforward excitatory and inhibitory synaptic plasticity, based on simple microscopic potentiation and depression processes, involving a unary variable, the synaptic weight m_H.

6. Conclusion

In this paper we have indicated how one can formulate and analyze actions for a simple network of excitatory cells with an input coupled to the network via an activity-dependent modifiable synapse. This action allows, in principal, the computation of statistical moments of the fluctuating dynamics of the network. Previous work by Wilson & Cowan (1972) indicates that the mean-field dynamics of the network is bistable, and can generate a hysteresis loop. On coupling this dynamics to that of the modifiable feedforward synapse introduced, all the conditions for the achievement of SOC are present in the combined system. The simulation of the mean-field dynamics shows that SOC is indeed achieved. It only remains to simulate the behavior of the combined

system with intrinsic noise. Our prediction is that the fluctuations in the activity near the critical points of the system will exhibit the properties of directed percolation. This will be the subject of another paper.

Acknowledgments

Appendix

Appendix A.1. Expanding the weighting function

We approximate the convolution $w \star n$ as:

$$
\begin{aligned}
w \star n &= \int d^d x' w(\mathbf{x} - \mathbf{x}') n(\mathbf{x}', t) \\
&\simeq (w_0 + \frac{1}{2!} w_2 \nabla^2 + \cdots) n(\mathbf{x}', t) \\
&\equiv Ln
\end{aligned}
\tag{A.1}
$$

where

$$
w(\underline{x}) \to w(r) = b/\sigma^d e^{-r/\sigma}, \sigma = r_0.
\tag{A.2}
$$

It follows that

$$
\begin{aligned}
w_0 &= \int d^d x w(\mathbf{x}) = b \frac{d\Gamma(d)}{\Gamma(d/2 + 1)} \pi^{d/2}, \\
w_2 &= \int d^d x^2 w(\mathbf{x}) = b\sigma^2 \frac{d\Gamma(d+2)}{\Gamma(d/2 + 1)} \pi^{d/2}
\end{aligned}
\tag{A.3}
$$

In case $d = 3$, $w_0 = 8\pi b$, $w_2 = 12 w_0 \sigma^2$, whence $w_2/w_0 = 12\sigma^2$. This expansion of the weighting function $w(\mathbf{x})$ is known as the *moment* expansion.

References

Abarbanel H & Broznan J 1974 *Phys. Lett. B* **48**, 345–348.
Abarbanel H, Broznan J, Schwimmer A & Sugar R 1976 *Phys. Rev. D* **14**(2), 632–646.
Amati D, Marchesini G, Ciafoloni M & Parisi G 1976 *Nuclear Physics B* **114**, 483–504.
Bak P, Tang C & Wiesenfeld K 1988 *Physical Review A* **38**(1), 364–374.
Beggs J & Plenz D 2003 *J. Neurosci.* **23**(35), 11167–11177.
Buice M A & Cowan J D 2007 *Physical Review E* **75**, 051919.
Buice M & Cowan J 2009 *Prog. Biophys. Theor. Biol.* **99**(2,3), 53–86.
Burns B D 1951 *J. Physiol.* **112**, 156–175.
Doi M 1976 *J. Phys. A: Math Gen.* **9**(9), 1465–1477.
Gil L & Sornette D 1996 *Physical Review Letters* **76**(21), 3991–3994.
Glauber R 1963 *Phys. Rev. Lett.* **10**(3), 84–86.
Lampl I, Reichova I & Ferster D 1999 *Neuron* **22**, 361–374.

Levina A, Herrmann J & Geisel T 2007 *Nature Physics* **3**(12), 857–860.

Millman D, Mihalas S, Kirkwood A & Niebur E 2010 *Nature Physics* **6**(10), 801–805.

Nauhaus I, Busse L, Carandini M & Ringach D 2009 *Nature Neurosci.* **12**(1), 70–76.

Peliti L 1985 *J. Physique* **46**, 1469–1483.

Schrödinger E 1926 *Naturwissenschaften* **14**, 664–666.

Stevens C 1989 *Neural Comp.* **1**, 473–479.

Van Kampen N 1981 *Stochastic Processes in Physics and Chemistry* North Holland.

Vogels T, Sprekeler H, Zenke F, Clopath C & Gerstner W 2011 *Science* **334**(6062), 664–666.

Wilson H & Cowan J 1972 *Biophys. J.* **12**, 1–22.

Wilson H R & Cowan J D 1973 *Kybernetik* **13**, 55–80.